Not quite a myriad of gene patents

Gregory D Graff, Devon Phillips, Zhen Lei, Sooyoung Oh, Carol Nottenburg & Philip G Pardey

A new study assesses the impact of recent US Supreme Court rulings on the changing landscape of US patents claiming nucleic acids.

Genetic innovations have long been jealously guarded—since well before Thomas Jefferson famously risked his life smuggling rice seed from the Piedmont region of Italy[1]. Since the early 1980s the United States Patent and Trademark Office (USPTO) has issued patents on inventions involving isolated DNA and RNA molecules with unique functionalities in living systems based on the sequence of nucleotides that make up the molecule[2], providing perhaps a better option to guard genetic innovations than 18th century Piedmont's threat of death.

However, patents claiming DNA molecules with nucleotide sequences that correspond directly to coding regions from the genomes of natural organisms, often called gene patents, have for decades elicited at least three fundamentally different kinds of objections[3–8]. First, there is a range of essentially moral arguments. Should patents be allowed to claim polynucleotide molecules if the sequences of those molecules—arguably their most useful characteristic—are products of nature, or evolution, rather than the fruits of human intellect? Also, under certain conditions, these intellectual property rights may conflict with other declared fundamental human rights, such as the right to know and use knowledge of one's own body, including one's own genetic sequences, the right to essential health care, or even reproductive rights. Second, there are utilitarian or pragmatic arguments. Do gene patents serve the intended purpose of stimulating innovation? Or, do gene patents, on balance, actually hinder innovation—or its utilization? Third, there are realist arguments, or what might be called public choice or political economic arguments between interest groups with differing economic stakes in the protection of genetic innovations, such as patients versus drug companies, farmers versus seed companies, generics manufacturers versus brand owners, technology-importing countries versus technology-exporting countries, and so on.

Motivated by recent and ongoing challenges to the patentability of human genes, we have undertaken a detailed, quantitative reassessment of the history of US patents in the technical area of genetics. The changing number of gene patents, subject matter orientations, and assignee status all hold implications for the consequences of legal decisions on the patentability of genes by the US Supreme Court.

The *Myriad* case

In the United States, the controversy over gene patents has recently culminated in the *Association of Molecular Pathology v. USPTO and Myriad Genetics* (often called the *Myriad* case)[9,10]. The fundamental question put before the US Supreme Court in this case is simply, "Are human genes patentable?" Specifically, the case challenges the validity of patent claims covering isolated DNA molecules that correspond to natural, albeit mutated, sequences of human *BRCA1* and *BRCA2* genes as invented compositions of matter. The case has also concerned claims to various methods of using those isolated DNA compositions to test women for genetic predisposition to breast cancer.

Initially, the lawsuit was filed in May 2009 with the US District Court of Southern New York. Judge Robert Sweet ruled in March 2010 that the isolated DNA molecules claimed as compositions of matter in the Myriad patents were not eligible subject matter, nor were the methods claimed for comparing gene sequences from patients against the sequences of those isolated DNA molecules to assess cancer risk[11]. The case was appealed to the US Court of Appeals for the Federal Circuit, which in July 2011 reversed in part the decision of the lower court, instead deciding that isolated DNA compositions are patentable[12]. A petition for *certiorari* was filed, appealing the case to the US Supreme Court.

In February 2012, after ruling in a different case, *Mayo Collaborative Services v. Prometheus Laboratories Inc.*[13], the Supreme Court remanded the *Myriad* case[14], sending it back to the Federal Circuit, from whence it had come, for reconsideration in light of its "law of nature" ruling in the *Mayo* case. The Supreme Court had ruled that Prometheus Laboratories' patent claims to a method for determining personalized doses of a drug—by monitoring natural levels of a metabolite of that drug in the patient's blood—were not patent eligible, on the grounds that the method simply involved the observation of an existing natural relationship without actually creating or inventing something new.

In August 2012, the Federal Circuit issued its second decision in the *Myriad* case[9]. Following the Supreme Court's guidance in *Mayo*, the Federal Circuit rejected the Myriad patents' claims to methods of comparing gene sequences from patients against the sequences of its isolated molecules; however, on the question of claims to those isolated DNA molecules

Gregory D. Graff is in the Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, Colorado, USA; Devon Phillips and Philip G. Pardey are in the Department of Applied Economics, University of Minnesota, St. Paul, Minnesota, USA; Zhen Lei is in the Department of Energy and Mineral Engineering and EMS Energy Institute, Pennsylvania State University, University Park, Pennsylvania, USA; Sooyoung Oh is in the Department of Computer Science and Engineering, Pennsylvania State University, University Park, Pennsylvania, USA; and Carol Nottenburg is at Cougar Patent Law, Renton, Washington, USA. Gregory D. Graff, Zhen Lei, Carol Nottenburg and Philip G. Pardey are also at the International Science and Technology Practice and Policy (InSTePP) Center, Department of Applied Economics, University of Minnesota, St. Paul, Minnesota, USA.
e-mail: gregory.graff@colostate.edu

as compositions of matter, the Federal Circuit reaffirmed its earlier decision that, to the extent they are new and different compositions of matter, "markedly different" in characteristics from what exists in nature in chromosomal form, they are eligible subject matter. Because in the eyes of the plaintiffs this decision did not resolve the fundamental disagreement, a petition for *certiorari* was once again filed at the Supreme Court.

In November 2012, the Supreme Court agreed again to hear the *Myriad* case, basically to resolve the question, "Are human genes patentable?", raising expectations that the Court would overturn or modify the Federal Circuit decision regarding composition-of-matter claims to isolated DNA molecules and establish an exception to patentability for at least some class of compositions consisting of naturally occurring DNA sequences[10].

The number of patents potentially affected by this case is generally thought to be in the thousands[10]. The August 2012 opinion of the Federal Circuit[9] cites several different references giving disparate counts of US gene patents, including Rogers' identification of the first two gene patents granted in 1982 (ref. 15); Caulfield, Gold and Cho's 1995 count of 1,750 human gene patents[16]; and the National Research Council committee's 2005 estimate of some 40,000 nucleotide-related patents[7]. In addition, a 2005 study by Jensen and Murray[17] identified 4,270 patents that cite human genetic sequences in the claims—sequences estimated to represent ~20% of the human genome. A more recent 2012 analysis, by Schauinger, identified 4,977 patents that claim human genetic sequences[18].

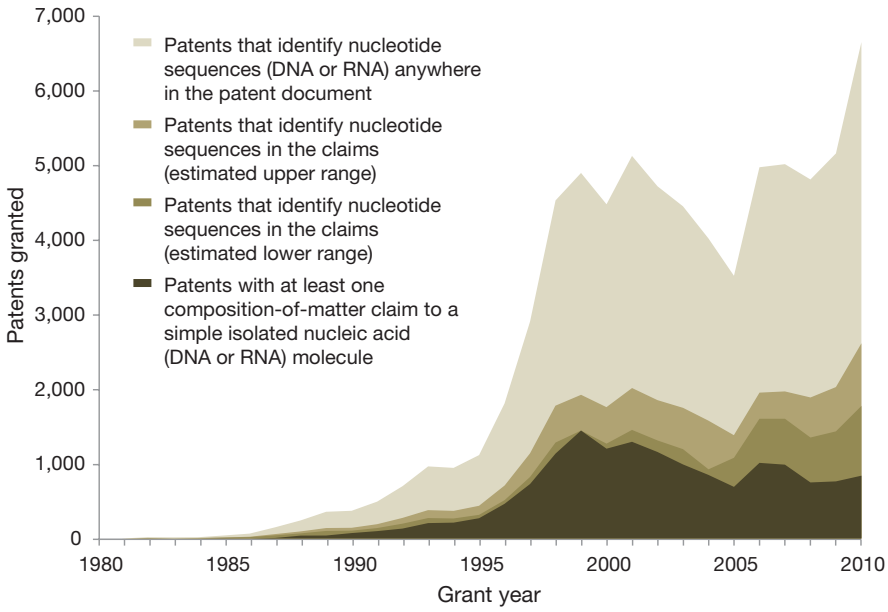

Figure 1 The shifting structure of US patents referring to and claiming nucleotide sequences. The darker categories are those more likely at risk if the "product of nature" exception to patent eligibility is extended to isolated nucleic acid molecules with naturally occurring sequences.

Types of claims most likely affected

The number of US patents affected by the *Myriad* case hinges on how the Supreme Court's decision affects various types of patent claims. The "claims" are the legal heart of a patent, written as a list of explicit statements summarizing exactly what technologies the patent protects. Patent claims are limited by law to one of four basic formats; a claim can be made (i) to a new method or process, (ii) to a new machine, (iii) to a new article of manufacture or (iv) to a new composition of matter. It is not uncommon for nucleotide sequences to show up in the text of a patent without actually being mentioned in or made subject to the claims of that patent. Moreover, not all patents that do recite nucleotide sequences in the claims are equally likely to be affected by the challenge raised in the *Myriad* case.

Claims can be made to methods of using a nucleic acid molecule (as a physical object) or its sequence (as abstract information), including methods of purifying molecules with that sequence and methods of identifying or diagnosing the presence of that sequence. Methods of simply comparing two genetic sequences with one another (as abstract information) have already been invalidated by the Federal Circuit's August 2012 decision in the *Myriad* case. With that settled, other types of method claims, to more complex uses of DNA, are no longer a focus in the case before the Supreme Court.

Claims can also be made to nucleic acid molecules as compositions of matter, and it is claims of this type that are the primary focus of the Supreme Court in the *Myriad* case. Still, not all composition-of-matter claims to DNA are considered equally problematic. Determining which ones are most likely to be affected, again, hinges on how the Supreme Court interprets the question, "Are human genes patentable?"

First is the issue of what exactly constitutes a "gene." Fundamentally, this is a question of when and under what conditions a DNA molecule—whether it codes for a protein or not—can be considered a "product of nature." What degree of transformation is sufficient to turn a segment of chromosomal DNA into a patentable invention? Although the isolation of a DNA molecule from the chromosome is arguably a transformation from what is found in nature, it is not a large transformation. Yet, there are further possible transformations of a segment of DNA, ranging from minor base-pair changes, to more extensive substitutions and deletions, to recombinations with other genetic elements. Where the line of demarcation is drawn by the Supreme Court between the "natural" and the "invented" determines to a great extent which patent claims are affected, and how.

A crucial distinction was made in the Federal Circuit's deliberations of the *Myriad* case between simple isolated DNA molecules with single natural sequences and engineered DNA molecules. Although it is possible for engineered DNA molecules to contain regions that correspond to naturally occurring genetic sequences, all three judges of the Federal Circuit appear to agree that, to the extent that changes are introduced in engineered constructs, they are more likely to be considered "markedly different" from what exists in nature and less likely to be considered "products of nature"[9]. Thus, some sort of demarcation between simple isolated nucleic acid molecules and complex engineered or recombinant nucleic acid molecules is essential in determining the extent to which different composition-of-matter claims are acceptable. Composition-of-matter claims to simple isolated DNA molecules are characterized by a single nucleotide sequence that is not situated within a combination of other genetic components or within a specified biological context. As a result, such claims tend to be quite broad. And, it is this last subset—composition-of-matter claims to simple, isolated DNA molecules with naturally occurring nucleotide sequences, copied or isolated in their entirety from somewhere in a genome—that most closely fit the idea of a "gene" as challenged in the legal arguments of the *Myriad* case.

Second is the issue of what is meant by "human." This is complicated by the extent to

which the human genome shares sequences in common with the genomes of other species, and by the fact that many patents already claim DNA isolated from such related species. It is not clear, going into the case, whether it really is the intent of the Supreme Court to single out *Homo sapiens* as a species and consider the patentability of only human genes. This implies the possibility that the Court might determine that only isolated DNA molecules with human genetic sequences are unpatentable products of nature, whereas isolated DNA with definitively non-human sequences remains patentable. Under such a biologically split decision, it would be necessary for the Court to provide clear rules for differentiating between (unpatentable) human sequences and (patentable) non-human sequences. This would be particularly important for those commercially important species that have at least segments of nucleotide sequences that are very close or even identical to human sequences, requiring a robust definition of what should be considered to constitute the human sequence and when and how it is to take precedence over such similar sequences isolated from other species. In other words, what are the tolerances around the human genome for what remains in the public domain? The impact of such a decision would presumably be to invalidate only composition-of-matter claims made to isolated human DNA, whereas composition claims to isolated DNA verifiably from other species would remain valid.

In practice, however, such delineation would be difficult to implement, particularly for patents already granted. Applicants are not required in the United States to identify the species from which a claimed sequence is isolated, and although in some patents the language of the claims text identifies the species from which a recited sequence originates, often it is not, and further investigation is required. At the same time, the main public genetics database, GenBank, systematically does not list the source species for nucleotide sequence accessions submitted from US patents. It is not immediately clear how the line should be drawn, in practice, between DNA from humans and DNA from other taxa[19].

Moreover, it is not clear whether it is even logical in principle to limit the question to *Homo sapiens*. If the fundamental legal question of patentability hinges on whether DNA molecules with naturally occurring sequences are "products of nature," it seems this principle should apply equally to any DNA molecule with a naturally occurring sequence, regardless of whether it originates from a human, or from another animal, plant, fungus, algae or microbe. This problem of delineation is particularly acute, again, with sequences that are highly conserved between humans and related species. If the Supreme Court decides that the product-of-nature argument applies only to humans, further legal challenges are likely to seek to extend the precedent to all isolated DNA from any species. If the Supreme Court decides in the current case that the principle applies more generally to species beyond *Homo sapiens*, then significantly more patent claims are potentially affected, including many involving the use of DNA in industrial applications well beyond the scope of human medicine, such as agriculture, food or biochemistry.

A third and final question is the extent to which the product-of-nature argument reaches beyond just DNA molecules and, at a minimum, also affects claims to isolated RNA molecules with naturally occurring nucleotide sequences. Indeed, as a practical matter, the language and the logic of composition-of-matter claims to RNA molecules, as typically written, can be indistinguishable from claims to DNA molecules, until one looks at the referenced sequences. And, again, it is not immediately clear how or why a line should be drawn between an isolated DNA molecule and an isolated RNA molecule, if both meet the condition of having naturally occurring sequences and thus being "products of nature"[20]. As a practical question, claims to both DNA and RNA need to be treated analytically as overlapping subsets within the larger set of claims to nucleic acid molecules.

Analysis

Given the range of different claim types, all of which can recite nucleotide sequences, the extent to which US patents contain these different types of claims is an empirical question that can be answered only by analyzing the language and the sequences cited within the claims of the relevant US patent literature.

The landscape of nucleotide-related patents. As a first step, to select the relevant literature, we drew upon a combination of sources[21–23] and query methods[24] (**Supplementary Note 1**) to find 72,052 US patents granted through the end of 2010 that in some way identify or make reference to nucleotide sequences (**Fig. 1**, top line).

Next, to establish a benchmark for how many of these patents actually make claims that involve nucleotide sequences, we infer, based upon proportions reported in the CAMBIA PatentLens sequence database[25,26], that somewhere between 30 and 39% of these 72,052 patents[27] are likely to contain or refer to nucleotide sequences in at least one of the claims of the patent (**Fig. 1**, middle lines). Thus, we infer that something between 21,870 and 28,410 granted US patents have a possibility of making a claim to a DNA or RNA molecule.

Claims analysis. To identify patents within this landscape that specifically contain composition-of-matter claims to simple isolated nucleic acids, it was necessary to shift our frame of analysis to the claims as listed in the patent texts. From the 72,052

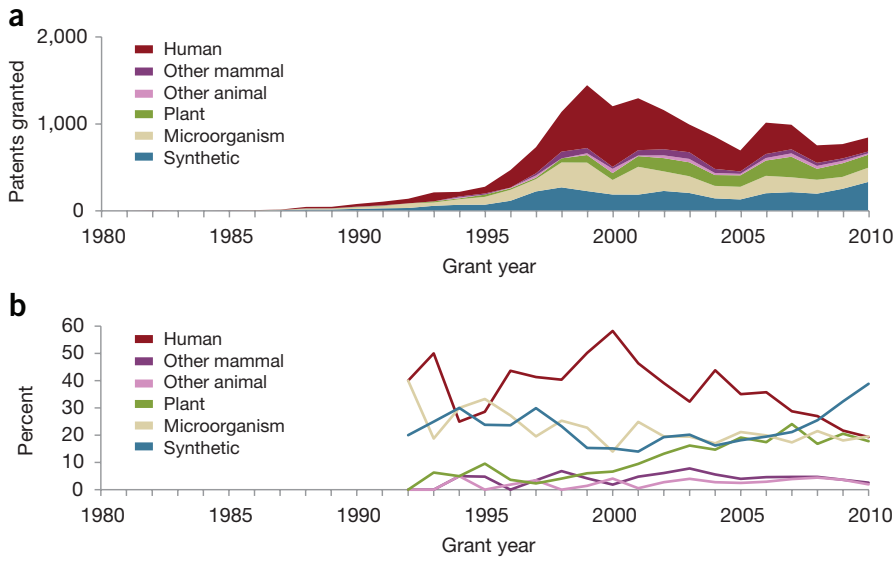

Figure 2 The primary sources of nucleotide sequences in the 15,359 gene patents (that is, patents with at least one composition-of-matter claim to a simple DNA or RNA molecule) granted in the United States. (**a**) Annual counts. (**b**) Annual shares.

nucleotide-related patents, we assessed all 213,128 independent claims (an average of 2.96 independent claims per patent). Of these, we found 82,952 (39%) were method claims, and were therefore set aside.

Linguistic algorithms were then used to analyze and categorize the remaining 130,176 (non-method) independent claims as to whether each is a composition-of-matter claim to simple nucleic acid molecules or not. The categorization results were reviewed by experts to resolve uncertain claims and to evaluate accuracy. From this analysis, 36,571 were recognized as composition-of-matter claims to simple isolated nucleic acid molecules[28] (**Supplementary Notes 1** and **2**).

Then, shifting the frame of analysis back to the level of patents, we found that these 36,571 composition claims to simple nucleic acids are contained within 15,359 of the 72,052 nucleotide-related patents (at an average of 2.38 per patent). It is this set of 15,359 patents containing at least one simple composition-of-matter claim to an isolated nucleic acid molecule that most closely fits the definition of "gene" patent as challenged in the *Myriad* case (**Fig. 1**, lowest line).

Source species analysis. We then turn to publicly available databases that designate the source or species of origin of listed nucleotide sequences[22,26]. From these, we estimate that 5,936 (39%) of these 15,359 gene patents primarily involve human sequences; 1,056 (7%) primarily involve other mammalian or animal sequences; 1,847 (12%) involve plant sequences; 3,228 (21%) involve microbial sequences; and 3,292 (21%) primarily involve simple sequences designated as "synthetic" or "artificial" (**Fig. 2** and **Supplementary Notes 1** and **2**).

The annual number of gene patents granted in the United States peaked in 1999 and has been in overall decline ever since. Since 2005, the annual grants of patents with claims to simple nucleotide molecules with naturally occurring sequences has stabilized around an average of 623 per year, roughly evenly split between human, plant and microbial sequences (**Fig. 2a**). The estimated share of all gene patents with human sequences peaked in 2000, at 58%, but by 2010 had fallen to 19% (**Fig. 2b**). The estimated share with synthetic sequences reached a low of 14% in 2001, but by 2010 had risen to almost 40% (**Fig. 2b**).

Although there are some questions about what constitutes a synthetic sequence, most are presumably altered from naturally occurring sequences. Thus, if we omit patents that primarily involve synthetic sequences as likely to be "markedly different" from naturally occurring DNA or RNA, we estimate that 11,868 US patents have been granted with composition-of-matter claims to simple nucleic acid molecules with naturally occurring sequences attributed to a specific species. Of these, 5,936 (precisely 50%) are attributed to humans.

Finally, we estimate that 8,703 (or 73%) of the 11,868 granted US gene patents with naturally occurring sequences are currently still in force (**Fig. 3**)[29]. Of these we estimate that 3,535 (41%) involve human sequences, 802 (9%) involve sequences of other mammalian and animal species, 1,693 (19%) involve plant sequences and 2,053 (24%) involve microbial sequences. Thus, 12% of the total 72,052 nucleotide-related patents found in the US patent literature have claims at risk of invalidation if the US Supreme Court overturns the Federal Circuit's decision. Based upon the species involved, it appears that less than half of these patents with at-risk claims protect applications in human medicine, whereas more than half protect applications in other fields of industry, such as veterinary medicine, crop agriculture, food and beverage manufacturing, industrial enzymes or bioenergy.

Patent assignee analysis. In considering the potential scope of commercial consequences of an overturn in the *Myriad* case, we next identify who owns the patents that are most at risk. The initial assignment of a patent provides an indication of who probably conducted the R&D and who may still have an economic stake in the patent rights. Of the 15,359 patents granted in the United States that contain composition-of-matter claims to simple nucleic acid molecules, 65% have private-sector (business) assignees only, whereas 24% have public-sector (government, university, nonprofit) assignees only. About 9% have a combination of public- and private-sector assignees (**Fig. 4**)[30]. These shares have been relatively stable over time (**Fig. 4c**) and are consistent with results found in previous studies[31,32].

Table 1 shows the top 25 assignees of these 15,359 US patents. These include large corporations, smaller biotech companies and a number of public-sector institutions. **Table 1** also shows estimated shares of source species characterizations for the gene patent portfolios of each of these top 25 assignees to highlight their diversity in terms of the market sectors of likely application. Many are pharmaceutical or biomedical companies, with portfolios primarily involving human genes. Several are large agricultural technology firms, with portfolios primarily focused on plant and microbial genes. In fact, DuPont, with its seed subsidiary, Pioneer Hi-Bred, focused on agricultural applications, is the single largest holder of US gene patents. Other firms with large agricultural genetics portfolios at risk include Monsanto, BASF and Bayer. The large public assignees have portfolios more balanced across species categories, as might be expected from institutions engaged in a wider array of research. Other assignees found in the sample, beyond the top 25, cover a wide range of commercial applications beyond human medicine—including agriculture, food and beverage manufacturing, biochemicals and enzyme manufacturing, and bioenergy.

Discussion

Several key trends and structural changes stand out in this landscape of US gene patents. Perhaps most striking—and most vividly remembered in the public imagination—was the exponential growth in all categories of nucleotide-related patenting from the mid-1980s up through 1998 or 1999 (**Fig. 1**). This period could be considered the early "homesteading" phase of genetics patenting, coinciding as it did with the initial sequencing of the human genome[33] and genomes of other important species.

After grant rates stabilized in 1999 to 2000, they then proceeded to decline through 2005, again in all categories of nucleotide-related patents (with the exception, perhaps, of plant genes) (**Figs. 1** and **2a**). A number of possible explanations could account for this general downturn. It coincides with the decline in venture capital available following the crash of the tech bubble[34]. It is also possible that the sheer volume of genetic sequences published in the scientific and patent literatures and posted to public sequence databases during the preceding boom years may have created so much prior art as to preclude a range of immediate follow-on patenting opportunities. Additionally, this downturn also coincided with the introduction of the USPTO Utility Examination Guidelines, first as interim guidelines in 1999 (ref. 35) and then in final form in 2001 (ref. 36), which significantly raised the examination standards for the claimed uses of genetic sequences.

We also see that a divergence began in 1999 between those patents that make composition claims to simple DNA or RNA molecules and those that make other kinds of claims involving nucleotide sequences, such as methods claims or composition claims to complex genetic constructs (**Fig. 1**). As the former began decreasing, the latter began to grow, particularly after 2005. By 2010, the number of US patents granted with composition-of-matter claims to simple isolated DNA or RNA molecules (estimated at 838 in 2010) was less than half the

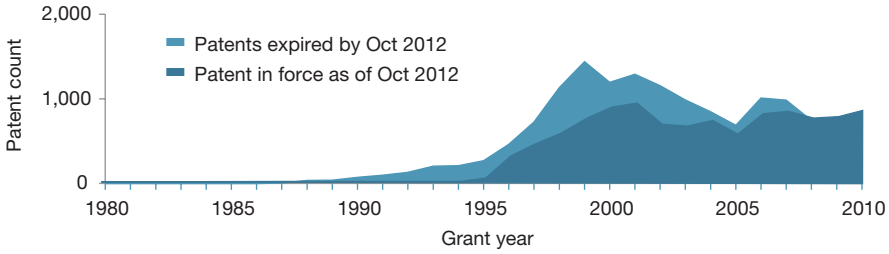

Figure 3 Current legal status of the 15,359 gene patents granted in the United States, by grant year.

total number of patents estimated to reference nucleotide sequences in the claims (estimated to be between 1,768 and 2,603 in 2010). This divergence may reflect the economic realization by companies, such as Celera Genomics and Incyte Pharmaceuticals, that it was proving difficult to profit from raw genetic sequences[37]. Indeed, Incyte has abandoned an unusually large share of its human gene patents (**Table 1**). The divergence is also consistent with the more stringent standards of the USPTO Utility Examination Guidelines, first signaled in 1999.

Finally, patents with human sequences, as a share of all gene patents, have declined from a high of 58% in 2000 to just 19% in 2010, the same share as plant sequences (**Fig. 2b**). In fact, by 2010, the largest share of gene patents were those designated as synthetic sequences, having grown over the same time period from an estimated 14% to almost 40% of all patents that claim simple nucleic acid molecules.

Based on analysis of claims language, we estimate the range of patents likely to be affected if the US Supreme Court extends a product-of-nature exception to isolated nucleic acid molecules with naturally occurring sequences. Beginning with the 72,052 granted US patents that can be associated with nucleic acid sequences, we found 8,073 granted US patents currently in force that contain the type of composition-of-matter claims—to simple nucleic acid molecules with natural sequences—that are most likely to be invalidated by a reversal of the *Myriad* case.

The question posed in the *Myriad* case focuses on human genetic sequences. Previously published studies also create the impression of pervasive patenting of the human genome[17,18]. We find that, of these 8,073 patents with composition-of-matter claims to simple nucleic acid molecules with natural sequences, 3,535 (41%) involve human genetic sequences whereas the other 4,538 (59%) involve sequences from other taxa, including animals, plants and microbes. Based upon the legal logic of the "product of nature" principle as well as the biological logic of the conservation of genetic sequences across species, we reason that the only workable product-of-nature exception would necessarily be broader in scope than just human genes. Therefore, an overturn by the Supreme Court will affect claims of patents not only on human genetic diagnostics and therapeutics but also on a wide range of other genetic technologies in other industries, particularly in agriculture, based upon our analysis of top assignee portfolios.

But, the demarcations used in this analysis to identify simple nucleotide sequences are not necessarily definitive. There are many patents with claims to complex or to simple synthetic DNA molecules for which validity may be difficult to resolve. How closely could their sequences match natural sequences? What, indeed, constitutes a natural sequence? What about single-site mutations or polymorphisms? At what point do they become "markedly different" from natural sequences? What if a variant synthesized in the laboratory today turns out to correspond to a natural sequence discovered at some point in the future? Such questions will likely perpetuate the controversies already at the heart of the *Myriad* case for years to come.

In the longer run, a judicial exception to the patentability of isolated nucleic acid molecules with naturally occurring sequences would likely only accelerate trends that are already very much evident in the data. More than a decade ago, applicants began to move away from drafting claims to simple isolated DNA molecules and from those with explicitly human sequences. Although the overall number of US patents that describe or ref-

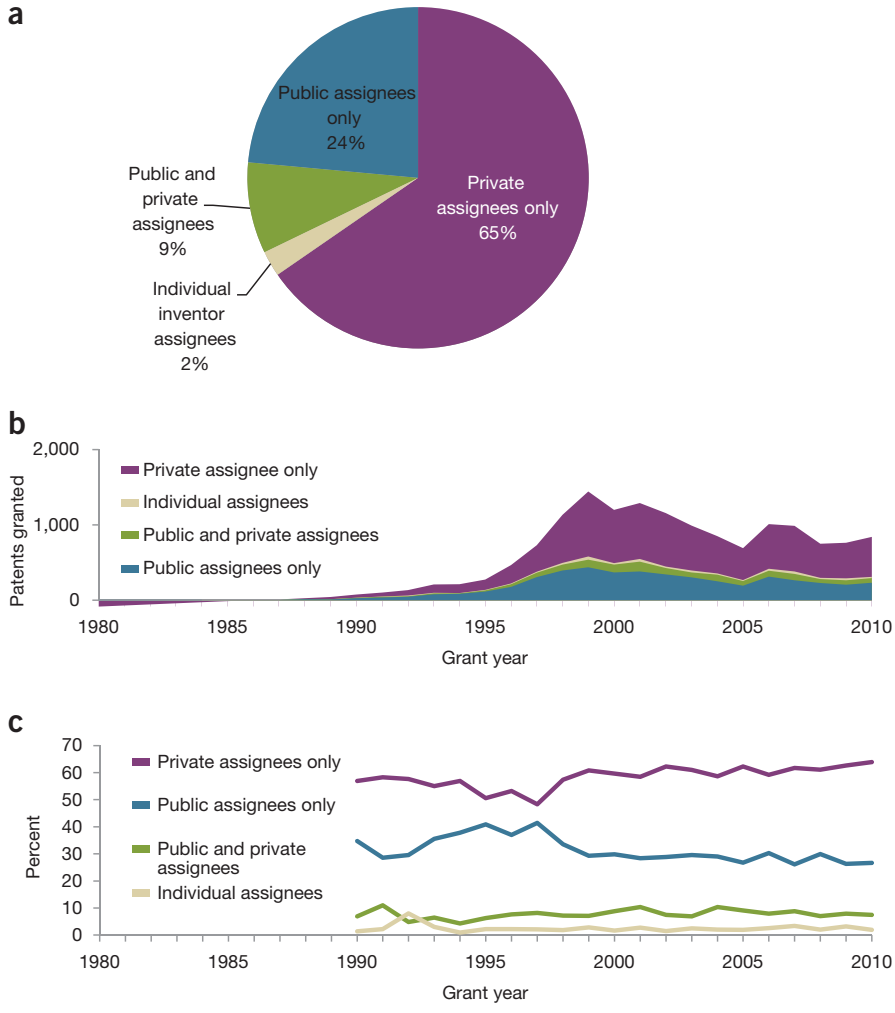

Figure 4 Shares of assignee types of the 15,359 gene patents granted in the United States. (**a**) Cumulative shares. (**b**) Annual counts. (**c**) Annual shares.

Table 1 Top 25 consolidated assignee portfolios of the 15,359 gene patents granted in the United States, by patent count

US "gene" patents					Assignee's shares by source of sequence (%)					
Assignee portfolio	Number granted	Number in force, as of Oct. 2012	Headquarters location	Assignee type	Human	Other mammal	Other animal	Plant	Micro-organism	Synthetic
DuPont/Pioneer Hi-Bred	764	655	USA	Private	0	0	3.4	68.4	17.5	10.7
GlaxoSmithKline	534	342	UK	Private	46.4	0.9	0	0	45.5	7.1
Roche	483	335	Switzerland	Private	67.5	5.2	0	0	8.4	18.8
Incyte	468	48	USA	Private	100.0	0	0	0	0	0
Pfizer	364	217	USA	Private	50.4	5.7	2.1	3.5	22.7	15.6
University of California	351	256	USA	Public	28.9	7.7	6.3	20.4	7.7	28.9
Isis Pharmaceuticals	350	318	USA	Private	55.7	0.5	0	0	0.5	43.4
Merck	338	224	USA	Private	42.4	9.8	6.1	0	10.6	31.1
Novartis	306	173	Switzerland	Private	24.5	0.0	0	6.4	37.2	31.9
US Dept. of Health and Human Services	302	219	USA	Public	36.4	3.6	1.8	0.9	14.5	42.7
Takeda Pharmaceuticals	263	174	Japan	Private	91.4	2.6	0	0	0	6.0
Monsanto	257	227	USA	Private	0.6	0.6	2.4	57.7	13.1	25.6
Bristol-Myers Squibb	243	156	USA	Private	50.8	3.2	0.8	0	12.7	32.5
Sanofi-aventis	242	180	France	Private	20.0	3.0	6.0	2.0	51.0	18.0
Amgen	239	160	USA	Private	81.3	3.8	0	0	0	15.0
Human Genome Sciences	237	148	USA	Private	82.8	0.0	0	0	0	17.2
Bayer	204	146	Germany	Private	22.2	0.0	4.4	30.0	26.7	16.7
BASF	158	141	Germany	Private	0.0	0.0	1.5	65.4	29.2	3.8
University of Washington	145	107	USA	Public	27.7	4.3	2.1	36.2	19.1	10.6
Life Technologies	140	57	USA	Private	60.0	2.5	10.0	0	10.0	17.5
Institut Pasteur	136	78	France	Public	4.0	0.0	2.0	0	56.0	38.0
Novo Nordisk	135	78	Denmark	Private	5.1	1.7	3.4	0	81.4	8.5
Johns Hopkins University	127	78	USA	Public	56.8	18.2	4.5	2.3	6.8	11.4
University of Texas	126	65	USA	Public	28.1	18.8	3.1	0	31.3	19.0
Eli Lilly	125	78	USA	Private	63.6	9.1	0	0	0	27.3

erence nucleotide sequences has continued to grow rapidly—with an estimated 6,619 granted in 2010 (**Fig. 1**, top line)—fewer and fewer of these—an estimated 512 granted in 2010—claim simple isolated molecules with naturally occurring sequences as compositions of matter (the nonsynthetic categories in **Fig. 2a**). Future patent applications that might have sought such claims—thus falling into this already diminishing category—could be readily redrafted to comply with new legal interpretations of subject matter eligibility. The very same sequences would likely still be the object of composition-of-matter claims, just claimed more obliquely, within the context of sufficiently complex, nonnative genetic constructs, or with enhancing changes to the sequence, enough to make it a "synthetic" or "artificial" sequence, rather than a "natural" sequence.

The outcome of the *Myriad* case, regardless, is thus likely to be less profound than either abolitionists or advocates seem to expect. In the end, any policy effects resulting from the *Myriad* case may prove to have been largely preempted by a combination of less dramatic changes in examination practices and patent law. We have hypothesized that the strengthening of utility requirements in 2001 by the USPTO was one cause of the observed slowdown in simple gene patents. The impacts of more recent changes—including the USPTO's limitation in 2007 to claiming just one sequence per patent under the "unity of invention" principle[38] and the Federal Circuit's *In re* Kubin ruling in 2009, which expanded obviousness requirements to include genetic analyses that can be considered "obvious to try"[39]—have likely not even begun to show up in the data analyzed here but will likely further reduce the already small annual number of new claims to simple isolated DNA molecules. Meanwhile, the patenting of engineered, "non-natural" genetic constructs, those most important to the various commercial applications of biotechnology, is likely to continue to surge ahead.

Note: Supplementary information is available in the online version of the paper (doi:10.1038/nbt.2568).

ACKNOWLEDGMENTS

The authors sincerely thank E. Hicks, K. Lee and C. Pratt for capable research assistance, and K. Silverstein for helpful advice. This research was supported by the US National Institutes of Health grant number 5 R01 HG004041-03.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Fowler, C. *Unnatural Selection* (Gordon and Breach, Switzerland, 1994).
2. For clarity of exposition, we seek throughout to maintain a semantic distinction between the term "nucleic acid molecule" as a physical object and the term "nucleotide sequence" as abstract information or code describing the structure of such a molecule.
3. Heller, M.A. & Eisenberg, R.S. *Science* **280**, 698–701 (1998).
4. Gold, E.R. *Nature* **18**, 1217–1218 (2000).
5. Nuffield Council on Bioethics. *The Ethics of Patenting DNA: a Discussion Paper* (2002) <http://www.nuffield-bioethics.org/patenting-dna>.
6. Resnik, D. *Owning the Genome: a Moral Analysis of DNA Patenting* (State University of New York Press, Albany, NY, 2004).
7. National Research Council. *Reaping the Benefits of Genomic and Proteomic Research: Intellectual Property Rights, Innovation, and Public Health* (National Academies Press, Washington, DC, 2006).
8. Holman, C.M. *Univ. Missouri Kansas City Law Rev.* **76**, 295–361 (2007).
9. *Association for Molecular Pathology v. USPTO*, No. 2010–1406 (Fed. Cir. Aug. 16, 2012).
10. Cook-Deegan, R. *Science* **338**, 745–747 (2012).
11. 702 F. Supp. 2d 181 (S.D.N.Y. 2010).
12. 653 F.3d 1329 (Fed. Cir. 2011).
13. 132 S. Ct. 1289 (2012).
14. 132 S. Ct. 1794 (2012).
15. Rogers, E.J. *J. Pat. Trademark Off. Soc.* **93**, 19–56 (2010).

16. Caulfield, T., Gold, E.R. & Cho, M. *Nat. Rev. Genet.* **1**, 227–231 (2000).
17. Jensen, K. & Murray, F. *Science* **310**, 239–240 (2005).
18. Schauinger, S. *The Human Gene Patent Report.* <http://hgpr.org/finalHGPR.pdf> (2012).
19. Moreover, some claims specify sequences within a certain percentage of homology or similarity to a reference sequence, which for more highly conserved genetic sequences, can, in theory, encompass multiple species' versions of that gene.
20. Moreover, if interpreted more broadly, the "product of nature" exception being argued could logically be extended to composition-of-matter claims to other biological molecules with naturally occurring sequences or structures, including proteins and antibodies, but logically even fatty acids, vitamins, etc.
21. National Center for Biotechnology Information (NCBI). GenBank Nucleotide database. <http://www.ncbi.nlm.nih.gov/nuccore/>
22. Benson, D.A. *et al. Nucleic Acids Res.* **33**, D34–D38 (2005).
23. Anonymous. Derwent World Patents Index (DWPI) CPI Manual Codes, Edn. 17, (Thomson Reuters, 2010). <http://ip-science.thomsonreuters.com/m/pdfs/cpi_manualcodes.pdf>
24. US patents in this analysis include all US patents referenced in NCBI's GenBank nucleotide sequence database, all US patents that responded to a full text query using all possible permutations of six-nucleotide (hexi-mer) sequences for both DNA and RNA as search terms, and all US patents identified by a select list of Derwent World Patent Index Manual Codes that refer to nucleotide sequences.
25. CAMBIA. PatentLens Sequence Search Facility. <http://www.patentlens.net/sequence/blast/blast.html>
26. Bacon, N. *et al.* (CAMBIA PatentLens, 2006). <http://www.patentlens.net/daisy/patentlens/2205.html>
27. Given the wide variety of ways in which a nucleotide sequence can be identified, such assessment is not straightforward. Bacon *et al.*[26] analyzed 31,572 US patents granted between 2002 and 2010, almost all of which were referenced by nucleotide accessions to the GenBank nucleotide database. In 12,240 (39%) of these patents at least one nucleotide sequence is found in at least one of the claims of the patent. In the remaining 19,332 (61%), nucleotide sequences were not found in any of the claims. Our range of estimates is based upon the assumption that any additional patents found by our search algorithms have the same or lower probability of containing nucleotide sequence references in the claims as those included in the CAMBIA PatentLens data.
28. The precision and recall rates of the machine learning algorithms used (that is, the true positives relative to classification by the algorithm and to classification by experts, respectively, for subsets of the training sample) was greater than 93%.
29. Expiration was either (a) projected from patent application and grant dates or (b) determined from the International Patent Documentation Center (INPADOC) Legal Status data of the European Patent Office (EPO) (http://www.epo.org/searching/subscription/raw/product-14-11.html), such as for failure to pay maintenance fees.
30. Interestingly, all three of the patents being challenged in the *Myriad* case (US patents 5,747,282, 5,837,492, and 5,693,473) involve both public and private sector assignees. See **Supplementary Notes 1** and **2**, appendix S1, for copies of the front pages of these three patents.
31. Graff, G.D. *et al. Nat. Biotechnol.* **21**, 989–995 (2003).
32. Pressman, L. *et al. Nat. Biotechnol.* **24**, 31–39 (2006).
33. Venter, J.C. *et al. Science* **291**, 1304–1351 (2001).
34. Goldfarb, B., Kirsch, D. & Miller, D. *J. Financial Econ.* **86**, 100–144 (2007).
35. *Interim Utility Examination Guidelines*, 64 Fed. Reg. 71440 (December 21, 1999).
36. *Utility Examination Guidelines*, 66 Fed. Reg. 1092 (January 5, 2001).
37. Kling, J. *EMBO Rep.* **6**, 1012–1014 (2005).
38. USPTO, 1316 O.G.13 (March 27, 2007).
39. *In re* Kubin, 561 F.3d 1351 (Fed. Cir. 2009).